

ERROR BOUNDS FOR CONSISTENT RECONSTRUCTION: RANDOM POLYTOPES AND COVERAGE PROCESSES

ALEXANDER M. POWELL AND J. TYLER WHITEHOUSE

ABSTRACT. Consistent reconstruction is a method for producing an estimate $\tilde{x} \in \mathbb{R}^d$ of a signal $x \in \mathbb{R}^d$ if one is given a collection of N noisy linear measurements $q_n = \langle x, \varphi_n \rangle + \epsilon_n$, $1 \leq n \leq N$, that have been corrupted by i.i.d. uniform noise $\{\epsilon_n\}_{n=1}^N$. We prove mean squared error bounds for consistent reconstruction when the measurement vectors $\{\varphi_n\}_{n=1}^N \subset \mathbb{R}^d$ are drawn independently at random from a suitable distribution on the unit-sphere \mathbb{S}^{d-1} . Our main results prove that the mean squared error (MSE) for consistent reconstruction is of the optimal order $\mathbb{E}\|x - \tilde{x}\|^2 \leq K\delta^2/N^2$ under general conditions on the measurement vectors. We also prove refined MSE bounds when the measurement vectors are i.i.d. uniformly distributed on the unit-sphere \mathbb{S}^{d-1} and, in particular, show that in this case the constant K is dominated by d^3 , the cube of the ambient dimension. The proofs involve an analysis of random polytopes using coverage processes on the sphere.

1. INTRODUCTION

We consider the problem of estimating an unknown signal $x \in \mathbb{R}^d$ from a collection of $N \geq d$ noisy linear measurements

$$q_n = \langle x, \varphi_n \rangle + \epsilon_n, \quad 1 \leq n \leq N, \quad (1.1)$$

where $\{\varphi_n\}_{n=1}^N \subset \mathbb{R}^d$ is a known spanning set for \mathbb{R}^d , and where the unknown noise $\{\epsilon_n\}_{n=1}^N$ has been independently drawn according to the uniform distribution on a known interval $[-\delta, \delta]$. *Consistent reconstruction* is a method for producing an estimate $\tilde{x} \in \mathbb{R}^d$ of x from the noisy measurements (1.1). Consistent reconstruction selects \tilde{x} as any solution to the linear feasibility problem

$$\forall 1 \leq n \leq N, \quad |\langle \tilde{x}, \varphi_n \rangle - q_n| \leq \delta. \quad (1.2)$$

In other words, (1.2) simply seeks an estimate \tilde{x} that is consistent with the knowledge that the noise is bounded in $[-\delta, \delta]$. Our main contribution in this work is to provide sharp bounds on the mean squared estimation error associated with consistent reconstruction and to quantify how accurately (1.2) recovers x from the measurements (1.1) as a function of the number of measurements N and the dimension d . In our analysis, $\{\varphi_n\}_{n=1}^N \subset \mathbb{R}^d$ will be i.i.d. random vectors drawn from a suitable distribution on the unit-sphere \mathbb{S}^{d-1} , and special attention will be given to the case when each φ_n is uniformly distributed on \mathbb{S}^{d-1} .

Consistent reconstruction has received particular attention in the signal processing literature as a method for recovering signals from quantized samples. Deterministic round-off errors that arise in quantization are frequently modeled using uniform noise. For example,

Date: April 9, 2013.

Key words and phrases. Consistent reconstruction, estimation with uniform noise.

the use of uniform noise models in quantization is typically justified by dithering, [14], or with high resolution asymptotics as the quantizer step-size approaches zero, [11]. Consistent reconstruction and its variants have been shown to be an effective method for signal recovery in memoryless scalar quantization (MSQ), [8, 19, 14, 3, 4], Sigma-Delta ($\Sigma\Delta$) quantization, [18], compressed sensing, [9], and finite rate of innovation sampling, [12]. A key point often observed in practice is that when compared with linear reconstruction, consistent reconstruction can reduce the mean squared reconstruction error by an extra multiplicative factor that scales inversely with the sampling rate.

The structure of uniformly distributed noise plays an essential role in our analysis of consistent reconstruction (1.2). It is useful to note that estimation with uniform noise in (1.1) falls outside of several classical approaches to estimation theory. For example, asymptotic normality theorems in maximum likelihood estimation (MLE) typically require suitable smoothness assumptions on the underlying noise distribution which uniform noise does not satisfy, e.g., [5]. Similarly, the multiparameter Cramer-Rao bound gives lower bounds for minimum variance unbiased estimation, but also requires suitable regularity on the noise distribution. Finally, recall that linear estimation commonly yields mean squared error (MSE) bounds of order $1/N$, and this is, for example, optimal for Gaussian noise with respect to the Cramer-Rao bound. However, when dealing with uniform noise, linear estimation is typically sub-optimal and it is possible to provide more accurate recovery than MSE of order $1/N$. We provide a rigorous analysis of the mean squared error in consistent reconstruction for the general estimation problem (1.1) with suitable random measurement vectors $\{\varphi_n\}_{n=1}^N \subset \mathbb{R}^d$, and we prove that the mean squared error is of the optimal order K/N^2 with precise control on the constant K . A sample consequence of this is that $N = \mathcal{O}(d^{3/2})$ random measurements will suffice to achieve highly accurate mean squared reconstruction error (compared to $\mathcal{O}(d^2)$ measurements with linear reconstruction).

Consistent reconstruction (1.2) has a simple geometrical interpretation. Since the bounded noise satisfies $|\epsilon_n| \leq \delta$, each noisy measurement q_n gives the information that the unknown true signal $x \in \mathbb{R}^d$ lies in the 2δ -thick slab

$$S_n = S_n(q_n, \varphi_n) = \{u \in \mathbb{R}^d : |\langle u, \varphi_n \rangle - q_n| \leq \delta\}. \quad (1.3)$$

Consequently, the consistency equations (1.2) are equivalent to requiring that $\tilde{x} \in \mathbb{R}^d$ lies in the *consistent reconstruction polytope* defined by

$$Q_N = \bigcap_{n=1}^N S_n. \quad (1.4)$$

The assumption that $\{\varphi_n\}_{n=1}^N$ spans \mathbb{R}^d ensures that Q_N is a compact set. Moreover, Q_N is almost surely a nondegenerate polytope with nonempty interior. Since the true signal x is always contained in the polytope Q_N , it is clear that the system (1.2) is feasible.

The main object of interest in this article will be the worst case error associated with consistent reconstruction. Recalling that (1.2) generally has infinitely many solutions, the worst case error may be defined as follows. If $\tilde{x} \in Q_N$ is any solution to the consistent

reconstruction system (1.2), then the error $(x - \tilde{x})$ lies in the following *error polytope*

$$P_N = \bigcap_{n=1}^N E_n \quad (1.5)$$

where

$$E_n = \{u \in \mathbb{R}^d : |\langle u, \varphi_n \rangle - \epsilon_n| \leq \delta\}. \quad (1.6)$$

The polytope P_N is obtained by translating Q_N to the origin by the vector x . In terms of the polytopes P_N and Q_N , the worst case error associated with consistent reconstruction can be defined as

$$W_N = \max\{\|u\| : u \in P_N\} = \max\{\|u - x\| : u \in Q_N\}. \quad (1.7)$$

Viewed geometrically, the worst case error W_N is precisely the radius of the smallest closed ball centered at 0 that contains the error polytope P_N .

Overview and main results. The main contribution of this article is to prove that the expected worst case error squared for consistent reconstruction is of the optimal order

$$\mathbb{E}|W_N|^2 \leq \frac{K\delta^2}{N^2}, \quad (1.8)$$

where the constant $K > 0$ depends on the distribution of the random vectors $\{\varphi_n\}_{n=1}^N$. Our first main result, Theorem 5.5, proves the mean squared error bound (1.8) under general assumptions on the i.i.d. unit-norm random vectors $\{\varphi_n\}_{n=1}^N \subset \mathbb{S}^{d-1}$. Our second main result, Theorem 6.1, proves a refined version of (1.8) when $\{\varphi_n\}_{n=1}^N \subset \mathbb{S}^{d-1}$ are i.i.d. random vectors that are uniformly distributed on \mathbb{S}^{d-1} and, in particular, shows that the constant K in (1.8) is dominated by d^3 , the cube of the ambient dimension.

The paper is organized as follows. Section 2 provides background on estimation with uniform noise, and also discusses preliminaries and notation such as surface measure on the sphere and geodesic ϵ -nets. Section 3 analyzes the size of the error polytope P_N in a fixed direction. Section 4 provides necessary background and results on coverage processes on the sphere which will be used in the proofs of our main theorems. Section 5 states and proves our first main theorem, Theorem 5.5, which shows that consistent reconstruction satisfies (1.8) under general assumptions. Section 6 states and proves our second main theorem, Theorem 6.1, which shows that for random vectors that are uniformly distributed on the unit-sphere consistent reconstruction satisfies (1.8) with a constant K that is dominated by d^3 .

2. BACKGROUND AND NOTATION

2.1. Estimation with uniform noise and consistent reconstruction. In this section we briefly recall some background on estimation with uniform noise and consistent reconstruction. This will help provide perspective on our main results.

Begin by recalling linear reconstruction. Let $N \geq d$. Suppose that $\{\varphi_n\}_{n=1}^N \subset \mathbb{R}^d$ spans \mathbb{R}^d and that $\{f_n\}_{n=1}^N \subset \mathbb{R}^d$ is any dual frame satisfying

$$\forall x \in \mathbb{R}^d, \quad x = \sum_{n=1}^N \langle x, \varphi_n \rangle f_n. \quad (2.1)$$

Suppose for the moment that $\{\epsilon_n\}_{n=1}^N$ are simply independent zero mean random variables with variance $\mathbb{E}|\epsilon_n|^2 = \sigma^2$, and that one wishes to estimate $x \in \mathbb{R}^d$ from the noisy measurements $q_n = \langle x, \varphi_n \rangle + \epsilon_n$, $1 \leq n \leq N$. Using the dual frame $\{f_n\}_{n=1}^N$ to linearly reconstruct $\tilde{x} \in \mathbb{R}^d$ by

$$\tilde{x} = \sum_{n=1}^N q_n f_n \quad (2.2)$$

yields the estimation error

$$\forall x \in \mathbb{R}^d, \quad \mathbb{E}\|x - \tilde{x}\|^2 = \sigma^2 \sum_{n=1}^N \|f_n\|^2. \quad (2.3)$$

If each φ_n is assumed to be unit-norm, $\|\varphi_n\| = 1$, then the MSE in (2.3) is bounded below by

$$\forall x \in \mathbb{R}^d, \quad \mathbb{E}\|x - \tilde{x}\|^2 \geq \frac{d^2 \sigma^2}{N}, \quad (2.4)$$

for example, see [7]. Moreover, if each $\|\varphi_n\| = 1$ then equality holds in (2.4) precisely when $\{\varphi_n\}_{n=1}^N$ is a unit-norm tight frame for \mathbb{R}^d and when $f_n = \frac{d}{N} \varphi_n$ is taken as the associated canonical dual frame, e.g., see [7]. The case of unit-norm tight frames yields the mean squared error $\mathbb{E}\|x - \tilde{x}\|^2 = d^2 \sigma^2 / N$.

It is natural to ask how much one can improve on the standard $1/N$ mean squared accuracy (2.4) if one uses nonlinear reconstruction. We are specifically interested in the case of uniform noise, and assume henceforth that $\{\epsilon_n\}_{n=1}^N$ are i.i.d. uniform random variables on $[-\delta, \delta]$.

We begin by mentioning a Bayesian lower bound due to Rangan and Goyal, [14]. Suppose that the vectors $\{\varphi_n\}_{n=1}^\infty \subset \mathbb{R}^d$ are unit-norm and that $x \in \mathbb{R}^d$ is an absolutely continuous random vector and that $q_n = \langle x, \varphi_n \rangle + \epsilon_n$. Let $\tilde{x}_N = \tilde{x}_N(\{q_n\}_{n=1}^N, \{\varphi_n\}_{n=1}^N)$ be any estimator which maps each observed input $(\{q_n\}_{n=1}^N, \{\varphi_n\}_{n=1}^N)$ to an estimate $\tilde{x}_N \in \mathbb{R}^d$ of the signal $x \in \mathbb{R}^d$. It was shown in [14] that the mean squared error is bounded below in the following manner:

$$\liminf_{N \rightarrow \infty} N^2 \mathbb{E}\|x - \tilde{x}_N\|^2 > 0. \quad (2.5)$$

Unlike (2.4) the expectation in (2.5) is taken over both a random signal x and the noise $\{\epsilon_n\}_{n=1}^N$. The lower bound (2.5) shows that for estimation in the setting of uniform noise one cannot expect MSE that is more accurate than $1/N^2$. Related lower bounds for quantization problems can be found in [8].

In [14], Rangan and Goyal proposed an estimation algorithm for (1.1) that achieves the optimal $1/N^2$ error rate. Their algorithm starts with an arbitrary $\tilde{x}_0 \in \mathbb{R}^d$, and iteratively produces estimates $\tilde{x}_n \in \mathbb{R}^d$ with the following soft-thresholding algorithm

$$\tilde{x}_n = \tilde{x}_{n-1} + \frac{\varphi_n T_\delta(q_n - \langle \tilde{x}_{n-1}, \varphi_n \rangle)}{\|\varphi_n\|^2}, \quad (2.6)$$

where the soft-thresholding function $T_\delta : \mathbb{R} \rightarrow \mathbb{R}$ is defined by

$$T_\delta(t) = \begin{cases} t - \delta, & \text{if } t > \delta, \\ 0, & \text{if } |t| \leq \delta, \\ t + \delta, & \text{if } t < -\delta. \end{cases} \quad (2.7)$$

The error analysis of the algorithm (2.6) in [14] assumed that $\{\varphi_n\}_{n=1}^N$ are independent identically distributed versions of a random vector φ satisfying the following condition

$$\exists \alpha > 0, \forall x \in \mathbb{S}^{d-1}, \quad \mathbb{E}|\langle x, \varphi \rangle| \geq \alpha > 0. \quad (2.8)$$

It was proven in [14] that if the i.i.d. random vectors $\{\varphi_n\}_{n=1}^N$ satisfy (2.8) then

$$\forall s < 1, \forall x \in \mathbb{R}^d, \quad \lim_{N \rightarrow \infty} N^{2s} \|x - \tilde{x}_N\|^2 = 0, \quad \text{almost surely.} \quad (2.9)$$

Moreover, the mean squared error was later proven to satisfy $\mathbb{E}\|x - \tilde{x}_N\|^2 \leq C/N^2$ for a suitable constant $C > 0$ in [13]. The algorithm (2.6) need not produce globally consistent estimates but instead employs local updates that can be sensitive to ordering issues.

Consistent reconstruction (1.2) provides a maximum likelihood estimate (MLE) for (1.1). Let $q = q(x)$ be the $N \times 1$ random vector that is parametrized by x and whose n th entry is $q_n = \langle x, \varphi_n \rangle + \epsilon_n$. The associated likelihood function is given by $\mathcal{L}(x|q) = \chi_{Q_N}(x)$, where χ_{Q_N} is the indicator function of the consistent reconstruction polytope Q_N in (1.4). Thus, the likelihood function $\mathcal{L}(x|q)$ is maximized precisely when x is a consistent estimate satisfying (1.2). Moreover, asymptotic normality results from MLE do not apply here since $\mathcal{L}(x|q)$ does not satisfy the smoothness assumptions that are typically needed, [5].

We conclude this section with the following simple one-dimensional example to provide intuition into the desired $1/N^2$ error rate for consistent reconstruction.

Example 2.1 (Consistent reconstruction in one dimension). Let $x \in \mathbb{R}$ and $q_n = x + \epsilon_n$, where $\{\epsilon_n\}_{n=1}^N$ are i.i.d. uniform random variables on $[-\delta, \delta]$. Consider the problem of estimating x from the noisy observations $\{q_n\}_{n=1}^N$. In this one-dimensional example, consistent reconstruction simply selects any estimate \tilde{x}_N that lies in the interval $I_N = [A_N, B_N]$ where

$$A_N = \max\{q_n - \delta : 1 \leq n \leq N\} \quad \text{and} \quad B_N = \min\{q_n + \delta : 1 \leq n \leq N\},$$

and the associated worst case estimation error is given by

$$w_N = \max\{|x - A_N|, |x - B_N|\}.$$

Elementary order statistics computations show that

$$\mathbb{E}|x - A_N|^2 = \mathbb{E}|x - B_N|^2 = \frac{8\delta^2}{(N+1)(N+2)}$$

and that the worst error w_N satisfies the following mean squared error bound

$$\mathbb{E}|w_N|^2 = \frac{14\delta^2}{(N+1)(N+2)}.$$

An important technical issue for analyzing (1.2) in \mathbb{R}^d will be that the geometry of the error polytope (1.5) becomes non-trivial in higher dimensions.

2.2. Preliminaries and notation. In this section we collect some necessary notation and background results concerning measure on the sphere and epsilon-nets.

We shall denote an open spherical cap on the unit-sphere $\mathbb{S}^{d-1} \subset \mathbb{R}^d$ with center $\varphi \in \mathbb{S}^{d-1}$ and angular radius $0 < \theta < \pi$ by

$$\text{Cap}(\varphi, \theta) = \{u \in \mathbb{S}^{d-1} : \langle u, \varphi \rangle > \cos \theta\}. \quad (2.10)$$

If $0 < \theta < \pi/2$ then the relative measure (normalized with respect to \mathbb{S}^{d-1}) of $\text{Cap}(\varphi, \theta)$ is given by, e.g., [2],

$$r_{d-1}(\theta) = \frac{\Gamma(\frac{d}{2})}{\sqrt{\pi}\Gamma(\frac{d-1}{2})} \int_0^\theta (\sin u)^{d-2} du. \quad (2.11)$$

It is useful to note, e.g., [21], that when $d \geq 2$ the constant

$$C_d = \frac{\Gamma(\frac{d}{2})}{\sqrt{\pi}\Gamma(\frac{d-1}{2})} \quad (2.12)$$

satisfies

$$\sqrt{1 - \frac{1}{d}} \sqrt{\frac{d-1}{2\pi}} \leq C_d \leq \sqrt{\frac{d-1}{2\pi}} \quad \text{and} \quad \lim_{d \rightarrow \infty} \frac{C_d}{\sqrt{d}} = \frac{1}{\sqrt{2\pi}}. \quad (2.13)$$

Fix any $x_0 \in \mathbb{S}^{d-1}$ and let the random vector $\varphi \in \mathbb{R}^{d-1}$ be uniformly distributed on the unit-sphere \mathbb{S}^{d-1} . Note that, by rotation invariance, the distribution of the random variable $Z = |\langle x_0, \varphi \rangle|$ does not depend on x_0 . The pdf of the random variable Z is given by, e.g., [17],

$$f_Z(z) = \begin{cases} 2C_d(1 - z^2)^{\frac{d-3}{2}}, & \text{if } z \in [0, 1], \\ 0, & \text{if } z \notin [0, 1]. \end{cases} \quad (2.14)$$

The geodesic distance between two points $x, y \in \mathbb{S}^{d-1}$ on the sphere will be denoted by

$$d(x, y) = \arccos(\langle x, y \rangle).$$

Given $\epsilon > 0$, we say that a set $\mathcal{N}_\epsilon \subset \mathbb{S}^{d-1}$ is a geodesic ϵ -net of \mathbb{S}^{d-1} if

$$\forall x \in \mathbb{S}^{d-1}, \exists z \in \mathcal{N}_\epsilon, \quad \text{such that} \quad d(x, z) \leq \epsilon.$$

A standard argument shows that if $d \geq 2$ then there exist geodesic ϵ -nets \mathcal{N}_ϵ of \mathbb{S}^{d-1} with cardinality satisfying

$$\#(\mathcal{N}_\epsilon) \leq \frac{1}{r_{d-1}(\epsilon/2)} \leq \frac{\sqrt{\pi}\Gamma(\frac{d-1}{2})}{\Gamma(\frac{d}{2})} \left(\frac{\pi}{2}\right)^{d-2} \frac{(d-1)}{(\epsilon/2)^{d-1}} \leq \left(\frac{8}{\epsilon}\right)^{d-1}. \quad (2.15)$$

For example, if \mathcal{N}_ϵ is a maximal ϵ -separated (with respect to geodesic distance) subset of \mathbb{S}^{d-1} then \mathcal{N}_ϵ is a geodesic ϵ -net for \mathbb{S}^{d-1} and satisfies (2.15).

3. ERROR POLYTOPE SIZE IN A FIXED DIRECTION

In this section we study the radial size of the error polytope P_N in a *fixed* direction. Given a unit-vector $\psi \in \mathbb{S}^{d-1}$, define the radial size of the error polytope in the direction ψ by

$$R_N(\psi) = \max\{r \geq 0 : r\psi \in P_N\}. \quad (3.1)$$

Note that $W_N \geq R_N$ since the worst case error (1.7) satisfies

$$W_N = \sup\{R_N(\psi) : \psi \in \mathbb{S}^{d-1}\}.$$

The next result follows immediately from Example 2.1 and provides a simple lower bound on $\mathbb{E}|R_N(\psi)|^2$ for a general choice of $\{\varphi_n\}_{n=1}^N \subset \mathbb{S}^{d-1}$.

Proposition 3.1. *Let $\{\varphi_n\}_{n=1}^N \subset \mathbb{S}^{d-1}$ be arbitrary and let $\psi \in \mathbb{S}^{d-1}$. Then the worst case error in the direction ψ for consistent reconstruction satisfies*

$$\forall N, \quad \mathbb{E}|R_N(\psi)|^2 \geq \frac{8\delta^2}{(N+1)(N+2)}.$$

Proof. Note that

$$R_N(\psi) = \min\{x_n : 1 \leq n \leq N\},$$

where

$$x_n = \begin{cases} (\epsilon_n + \delta)/|\langle \varphi_n, \psi \rangle|, & \text{if } \langle \varphi_n, \psi \rangle \geq 0, \\ (\delta - \epsilon_n)/|\langle \varphi_n, \psi \rangle|, & \text{if } \langle \varphi_n, \psi \rangle < 0. \end{cases}$$

Since $|\langle \varphi_n, \psi \rangle| \leq 1$ and since $(\epsilon_n + \delta)$ and $(\delta - \epsilon_n)$ are both uniformly distributed on $[0, 2\delta]$ it follows that $x_n \geq \xi_n$, where ξ_n is uniformly distributed on $[0, 2\delta]$. The proof now follows from Example 2.1. \square

In the remainder of this section, we study $R_N(\psi)$ in the case when the vectors $\{\varphi_n\}_{n=1}^N \subset \mathbb{R}^d$ used to define the error polytope P_N are i.i.d. uniformly distributed random vectors on the unit-sphere \mathbb{S}^{d-1} . In this case, rotation invariance implies that the distribution of $R_N = R_N(\psi)$ is independent of ψ . We will make use of the following lemma whose proof follows from similar steps as in Proposition 3.1.

Lemma 3.2. *Let ξ be a uniform random variable on $[0, 2\delta]$, and define the random variable $Z = |\langle e_0, \varphi \rangle|$, where $e_0 = (1, 0, 0, \dots, 0) \in \mathbb{R}^d$ and the random vector φ is uniformly distributed on unit-sphere \mathbb{S}^{d-1} . Let ξ and Z be independent and define the random variable $X = \xi/Z$. Let $\{X_n\}_{n=1}^N$ be N independent versions of the random variable X and consider the associated order statistic*

$$Y_N = \min\{X_n : 1 \leq n \leq N\}$$

The random variable R_N has the same distribution as the random variable Y_N .

Theorem 3.3. *Let $d \geq 2$. Suppose that $\{\varphi_n\}_{n=1}^N \subset \mathbb{S}^{d-1}$ are i.i.d. uniformly distributed random vectors on \mathbb{S}^{d-1} . Then for $N \geq 3$ the worst case error in the direction $\psi \in \mathbb{S}^{d-1}$ for consistent reconstruction satisfies*

$$\mathbb{E}|R_N(\psi)|^2 = \frac{2\delta^2(d-1)^2}{(C_d)^2(N+1)(N+2)} + 2\delta^2\alpha_{d,N}, \quad (3.2)$$

where C_d is as in (2.12) and

$$-\left(\frac{2C_d}{(d-1)}\right) \left(1 - \frac{C_d}{(d-1)}\right)^{N+1} \leq \alpha_{d,N} \leq 54(C_d)^2 \left(1 - \frac{2C_d}{(d-1)}\right)^N. \quad (3.3)$$

Note that by (2.13), $0 < (1 - \frac{C_d}{d-1}) < 1$ and $0 < (1 - \frac{2C_d}{d-1}) < 1$.

Proof. By Lemma 3.2 we need to estimate the following integral

$$\mathbb{E}|R_N(\psi)|^2 = \mathbb{E}|Y_N|^2 = 2 \int_0^\infty \lambda \Pr[Y_N > \lambda] d\lambda = 2 \int_0^\infty \lambda (\Pr[X > \lambda])^N d\lambda. \quad (3.4)$$

Step I. If $0 \leq \lambda \leq 2\delta$ then using (2.14) gives

$$\begin{aligned} \Pr[X > \lambda] &= \Pr[\xi > \lambda Z] = \int_0^1 \Pr[\xi > \lambda z] f_Z(z) dz \\ &= \int_0^1 \left(\frac{2\delta - \lambda z}{2\delta} \right) 2C_d (1 - z^2)^{\frac{d-3}{2}} dz \\ &= \left(1 - \frac{\lambda C_d}{\delta(d-1)} \right). \end{aligned} \quad (3.5)$$

A computation shows that

$$\begin{aligned} \int_0^{2\delta} \lambda (\Pr[X > \lambda])^N d\lambda &= \int_0^{2\delta} \lambda \left(1 - \frac{\lambda C_d}{\delta(d-1)} \right)^N d\lambda \\ &= \frac{\delta^2(d-1)^2}{(C_d)^2} \left(\frac{1}{(N+1)(N+2)} - \left(1 - \frac{2C_d}{(d-1)} \right)^{N+1} \left(\frac{1}{N+1} - \frac{(1 - \frac{2C_d}{(d-1)})}{N+2} \right) \right). \end{aligned} \quad (3.6)$$

Step II. If $\lambda \geq 2\delta$ then then (3.5) gives

$$\Pr[X > \lambda] \leq \Pr[X > 2\delta] \leq \left(1 - \frac{2C_d}{(d-1)} \right).$$

Thus

$$\begin{aligned} \int_{2\delta}^{6\pi\delta C_d} \lambda (\Pr[X > \lambda])^N d\lambda &\leq \int_{2\delta}^{6\pi\delta C_d} \lambda \left(1 - \frac{2C_d}{(d-1)} \right)^N d\lambda \\ &\leq 18\pi^2\delta^2(C_d)^2 \left(1 - \frac{2C_d}{(d-1)} \right)^N. \end{aligned} \quad (3.7)$$

Step III. If $\lambda \geq 2\delta$ then a computation using (2.14) and $d \geq 2$ shows that

$$\begin{aligned} \Pr[X > \lambda] &= 2C_d \int_0^{2\delta/\lambda} \left(\frac{2\delta - \lambda z}{2\delta} \right) (1 - z^2)^{\frac{d-3}{2}} dz \\ &\leq 2C_d \int_0^{2\delta/\lambda} (1 - z^2)^{-\frac{1}{2}} dz \\ &\leq \frac{2\pi\delta C_d}{\lambda}. \end{aligned} \quad (3.8)$$

If $d \geq 2$ then (2.13) along with $\Gamma(1/2) = 1/\sqrt{\pi}$ and $\Gamma(3/2) = \sqrt{\pi}/2$ implies that $\frac{1}{3} \leq (1 - \frac{2C_d}{d-1})$. Equation (3.8) implies that

$$\begin{aligned} \int_{6\pi\delta C_d}^{\infty} \lambda (\Pr[X > \lambda])^N d\lambda &\leq \int_{6\pi\delta C_d}^{\infty} \lambda \left(\frac{2\pi\delta C_d}{\lambda} \right)^N d\lambda \\ &= \left(\frac{1}{3} \right)^N \frac{36\pi^2\delta^2(C_d)^2}{(N-2)} \\ &\leq 36\pi^2\delta^2(C_d)^2 \left(1 - \frac{2C_d}{(d-1)} \right)^N. \end{aligned} \quad (3.9)$$

Combining (3.6), (3.7), (3.9) and (3.4) now yields the desired conclusion (3.2). \square

Corollary 3.4. *Suppose that $\{\varphi_n\}_{n=1}^N \subset \mathbb{S}^{d-1}$ are i.i.d. uniformly distributed random vectors on \mathbb{S}^{d-1} . Then*

$$\liminf_{N \rightarrow \infty} N^2 \mathbb{E}|W_N|^2 \geq \lim_{N \rightarrow \infty} N^2 \mathbb{E}|R_N|^2 = 2\delta^2 \left(\frac{d-1}{2C_d} \right)^2 \geq \pi\delta^2(d-1).$$

4. CONSISTENT RECONSTRUCTION AND COVERAGE PROCESSES

In this section we describe a useful connection between the worst case error W_N and coverage processes on the sphere \mathbb{S}^{d-1} . This relationship will play a central role in the proofs of our upper bounds on $\mathbb{E}|W_N|^2$ that appear in the subsequent sections.

4.1. Worst case error and coverage processes. The expected worst case error squared $\mathbb{E}|W_N|^2$ can be represented as

$$\mathbb{E}|W_N|^2 = 2 \int_0^\infty \lambda \Pr[W_N > \lambda] d\lambda. \quad (4.1)$$

So, a main step towards bounding $\mathbb{E}|W_N|^2$ is to bound the probability $\Pr[W_N > \lambda]$. The next lemma shows that bounding $\Pr[W_N > \lambda]$ can be reformulated as a coverage problem.

Lemma 4.1. *For each $\lambda > 0$ define the set*

$$B_n = B_n(\lambda) = \left\{ u \in \mathbb{S}^{d-1} : \langle u, \varphi_n \rangle > \frac{\epsilon_n + \delta}{\lambda} \text{ or } \langle u, \varphi_n \rangle < \frac{\epsilon_n - \delta}{\lambda} \right\}. \quad (4.2)$$

Then

$$\forall \lambda > 0, \quad \Pr[W_N \geq \lambda] = \Pr \left(\mathbb{S}^{d-1} \not\subset \bigcup_{n=1}^N B_n(\lambda) \right). \quad (4.3)$$

In particular,

$$\forall \lambda > 0, \quad \Pr[W_N > \lambda] \leq \Pr \left(\mathbb{S}^{d-1} \not\subset \bigcup_{n=1}^N B_n(\lambda) \right). \quad (4.4)$$

Proof. Let E_n be as in (1.6) and define the set

$$A_n = A_n(\lambda) = (\lambda\mathbb{S}^{d-1}) \setminus E_n = \{u \in \lambda\mathbb{S}^{d-1} : \langle u, \varphi_n \rangle > \epsilon_n + \delta \text{ or } \langle u, \varphi_n \rangle < \epsilon_n - \delta\}.$$

Observe that

$$W_N \geq \lambda \quad \text{if and only if} \quad \lambda\mathbb{S}^{d-1} \not\subset \bigcup_{n=1}^N A_n(\lambda). \quad (4.5)$$

It only remains to rescale (4.5) to the unit-sphere. Using the map $u \mapsto u/\lambda$ and the set $B_n(\lambda)$ defined by (4.2) one has that

$$\lambda\mathbb{S}^{d-1} \not\subset \bigcup_{n=1}^N A_n(\lambda) \quad \text{if and only if} \quad \mathbb{S}^{d-1} \not\subset \bigcup_{n=1}^N B_n(\lambda). \quad (4.6)$$

The proof of (4.3) now follows by combining (4.5) and (4.6). \square

We shall refer to the set $B_n(\lambda)$ as a bi-cap since it can be expressed as the union of two antipodal (possibly empty) open spherical caps

$$B_n(\lambda) = \text{Cap}(\varphi_n, \theta_n^+) \cup \text{Cap}(-\varphi_n, \theta_n^-), \quad (4.7)$$

where the angular radii θ_n^+ and θ_n^- are given by

$$\theta_n^+ = \begin{cases} \arccos\left(\frac{\delta + \epsilon_n}{\lambda}\right), & \text{if } \delta + \epsilon_n < \lambda, \\ 0, & \text{otherwise,} \end{cases} \quad (4.8)$$

and

$$\theta_n^- = \begin{cases} \arccos\left(\frac{\delta - \epsilon_n}{\lambda}\right), & \text{if } \delta - \epsilon_n < \lambda, \\ 0, & \text{otherwise.} \end{cases} \quad (4.9)$$

In particular, depending on the size of the parameters ϵ_n and λ , each set $B_n(\lambda)$ is either: (i) a union of two disjoint spherical caps with antipodal centers, or (ii) a single spherical cap, or (iii) the empty set.

4.2. Background on coverage processes. In our general analysis of consistent reconstruction, the coverage problem in Lemma 4.1 involves spherical caps with both random angular radii and random centers. Random coverage problems have a long and technical history, e.g., see [16, 2], but in high dimensions the literature is still limited when considering caps with both random center and random size. In fact, even in the case of constant sized caps with random centers on \mathbb{S}^{d-1} , bounds on coverage probabilities were only recently obtained in [2]. Some noteworthy results for randomly sized caps include [15] which contains exact results in dimension $d = 2$ with general distributions on the random arclengths, and [10] which contains asymptotic results (as the random cap size becomes small) on general manifolds.

In this section, we shall provide some necessary background on coverage processes in the case of spherical caps with angular radii of fixed size and random centers $\{\varphi_n\}_{n=1}^N$ that are uniformly distributed on the unit-sphere. The techniques and results that we will use later are especially influenced by [2] and [6].

For the remainder of this section let $\{\varphi_n\}_{n=1}^N \subset \mathbb{S}^{d-1} \subset \mathbb{R}^d$ be i.i.d. random vectors that are uniformly distributed on \mathbb{S}^{d-1} , and let $0 < \theta < \pi/2$ be fixed. Consider the following non-coverage probability

$$p(N, d-1, \theta) = \Pr \left(\mathbb{S}^{d-1} \not\subset \bigcup_{n=1}^N \text{Cap}(\varphi_n, \theta) \right). \quad (4.10)$$

The following theorem contains the best known bounds on $p(N, d-1, \theta)$ when $0 < \theta < \pi/2$, see Theorem 1.1 and Proposition 5.5 in [2]. The work in [2] is stated for coverage by closed caps but the following result remains true for open caps.

Theorem 4.2 (Bürgisser, Cucker, Lotz). *If $0 < \theta < \pi/2$ and $N \geq d \geq 2$ then*

$$p(N, d-1, \theta) \leq 2^{1-N} \sum_{k=0}^{d-1} \binom{N-1}{k} + \binom{N}{d} \left(\frac{d\sqrt{d-1}}{2^{d-1}} \right) F_{N,d-1}(\theta), \quad (4.11)$$

where

$$F_{N,d-1}(\theta) = \int_0^{\cos \theta} (1-t^2)^{((d-1)^2-2)/2} (1-r_{d-1}(\arccos t))^{N-d-2} dt$$

and $r_{d-1}(\arccos t)$ is defined using (2.11).

We briefly comment on why Theorem 4.2 holds when $p(N, d-1, \theta)$ is defined as in (4.10) using open spherical caps instead of closed spherical caps as in [2]. For this it suffices to note that if $0 < \alpha < \theta < \pi/2$ then

$$\Pr \left(\mathbb{S}^{d-1} \not\subset \bigcup_{n=1}^N \text{Cap}(\varphi_n, \theta) \right) \leq \Pr \left(\mathbb{S}^{d-1} \not\subset \bigcup_{n=1}^N \overline{\text{Cap}(\varphi_n, \alpha)} \right),$$

and that

$$\lim_{\alpha \rightarrow \theta} F_{N,d-1}(\alpha) = F_{N,d-1}(\theta).$$

We shall later need bounds on $p(N, d-1, \theta)$ when $\arccos(1/\sqrt{d}) \leq \theta < \pi/2$, i.e., when the cap height is less than $1/\sqrt{d}$.

Lemma 4.3. *If $N \geq \frac{2d}{\ln(12/11)} \approx (22.99)d$ and $\arccos(1/\sqrt{d}) \leq \theta < \pi/2$ then*

$$p(N, (d-1), \theta) \leq 2\sqrt{d} (13)^d \left(\frac{11}{12} \right)^{N/2}.$$

In particular, if $s > \frac{2\ln(13)}{\ln(12/11)} \approx 58.96$ then

$$\lim_{d \rightarrow \infty} p(sd, (d-1), \theta) = 0. \quad (4.12)$$

Proof. Hoeffding's inequality shows that if $(N-1) \geq 4(d-1)$ then

$$2^{1-N} \sum_{k=0}^{d-1} \binom{N-1}{k} \leq e^{1/8} e^{-N/8}. \quad (4.13)$$

By Lemma 2.1 of [1] we have that

$$\arccos(1/\sqrt{d}) \leq \theta < \pi/2 \implies \frac{1}{12} \leq r_{d-1}(\theta) \leq \frac{1}{2}.$$

Thus, for $\arccos(1/\sqrt{d}) \leq \theta < \pi/2$ we have

$$\begin{aligned} F_{N,d-1}(\theta) &\leq \int_0^{1/\sqrt{d}} (1-t^2)^{((d-1)^2-2)/2} (1-r_{d-1}(t))^{N-(d-1)-1} dt \\ &\leq \frac{1}{\sqrt{d}} \left(\frac{11}{12}\right)^{N-d}. \end{aligned} \quad (4.14)$$

Combining (4.11), (4.13), (4.14) gives

$$\begin{aligned} p(N, d-1, \theta) &\leq e^{1/8} e^{-N/8} + \binom{N}{d} \frac{d\sqrt{d-1}}{2^{d-1}} \frac{1}{\sqrt{d}} \left(\frac{11}{12}\right)^{N-d} \\ &\leq e^{1/8} e^{-N/8} + 2d \left(\frac{12}{22}\right)^d \left(\frac{11}{12}\right)^N \binom{N}{d}. \end{aligned} \quad (4.15)$$

We shall use the following bounds to further simplify (4.15). First, note that by Stirling's approximation

$$\binom{N}{d} \leq \frac{N^d}{d!} \leq \frac{1}{\sqrt{2\pi d}} \left(\frac{eN}{d}\right)^d. \quad (4.16)$$

Also, it follows from $\ln(x) \leq (x/c) + (\ln(c) - 1)$ that one has

$$\forall c > 0, \quad \left(\frac{eN}{d}\right)^d \leq c^d e^{eN/c} e^{-d}. \quad (4.17)$$

Applying (4.16) and (4.17) with $c = 2e/\ln(12/11)$ to (4.15) yields that if $N \geq \frac{2d}{\ln(12/11)}$ then

$$\begin{aligned} p(N, d-1, \theta) &\leq e^{1/8} e^{-N/8} + 2d \left(\frac{12}{22}\right)^d \left(\frac{11}{12}\right)^N \binom{N}{d} \\ &\leq e^{1/8} e^{-N/8} + \frac{2\sqrt{d}}{\sqrt{2\pi}} \left(\frac{12}{22}\right)^d \left(\frac{2}{\ln(12/11)}\right)^d \left(\frac{11}{12}\right)^{N/2} \\ &\leq e^{1/8} e^{-N/8} + \sqrt{d} (13)^d \left(\frac{11}{12}\right)^{N/2} \\ &\leq 2\sqrt{d} (13)^d \left(\frac{11}{12}\right)^{N/2}. \end{aligned}$$

This completes the proof. □

5. UPPER BOUNDS FOR GENERAL DISTRIBUTIONS

In this section we prove that consistent reconstruction achieves MSE of the optimal order $\mathbb{E}|W_N|^2 \lesssim 1/N^2$ under rather general conditions on the i.i.d. random measurement vectors $\{\varphi_n\}_{n=1}^N \subset \mathbb{R}^d$.

Our error bounds use the following admissibility condition. We assume that $\{\varphi_n\}_{n=1}^N \subset \mathbb{R}^d$ are independent identically distributed versions of a unit-norm random vector $\varphi \in \mathbb{S}^{d-1}$, and we further assume that there exist constants $\alpha \geq 1, s > 0$ such that

$$\forall 0 \leq t \leq 1, \forall x \in \mathbb{S}^{d-1}, \quad \Pr(|\langle x, \varphi \rangle| \leq t) \leq \alpha t^s. \quad (5.1)$$

Roughly speaking, the admissibility condition (5.1) ensures that the random vector φ cannot be too concentrated on any subspace of \mathbb{R}^d with positive codimension.

Example 5.1. If $d \geq 3$ and if the random vector φ is uniformly distributed on the unit-sphere \mathbb{S}^{d-1} then φ satisfies (5.1) with $s = 1$ and $\alpha = 2C_d$, where C_d is as in (2.12). This follows since $0 \leq f_Z(z) \leq 2C_d$ in (2.14). Similarly, when $d = 2$, a direct computation shows that if φ is uniformly distributed on the unit-circle \mathbb{S}^1 then φ satisfies (5.1) with $s = 1$ and $\alpha = 1$. This follows using $2C_2 = 2/\pi$ and $\arcsin(t) \leq (\pi/2)t$.

Example 5.2. Suppose that the unit-norm random vector $\varphi_1 \in \mathbb{S}^{d-1}$ satisfies (5.1) with $\alpha = \alpha_1$ and $s = s_1$. Suppose that the unit-norm random vector $\varphi_2 \in \mathbb{S}^{d-1}$ has the property that there exists $\beta > 0$ such that for every Borel subset $B \subset \mathbb{S}^{d-1}$ there holds $\Pr[\varphi_2 \in B] \leq \beta \Pr[\varphi_1 \in B]$. Then φ_2 satisfies (5.1) with $\alpha = \beta\alpha_1$ and $s = s_1$.

Example 5.3. If the random vector φ is uniformly distributed on an open subset of \mathbb{S}^{d-1} then φ satisfies (5.1). This follows by combining Examples 5.1 and 5.2.

Example 5.4. If the random vector φ has a point mass at $v \in \mathbb{R}^d$ with $\Pr[\varphi = v] > 0$ then φ does not satisfy (5.1). To see this, let $x_v \in \mathbb{S}^{d-1}$ be any vector that is orthogonal to v . Then $\Pr[\langle \varphi, x_v \rangle = 0] \geq \Pr[\varphi = v] > 0$ shows that (5.1) does not hold.

We are now ready to state and prove our first main theorem.

Theorem 5.5. *Suppose that $\{\varphi_n\}_{n=1}^N \subset \mathbb{S}^{d-1}$ are i.i.d. versions of a unit-norm random vector φ that satisfies (5.1). For all $N \geq (d+2)/s$ the expected worst case error squared satisfies*

$$\mathbb{E}|W_N|^2 \leq \frac{10^5 \delta^2 d^2 (2\alpha)^{2/s} \ln^2(16(2\alpha)^{1/s})}{(N+1)(N+2)} + \delta^2 32^{d+1} (2\alpha)^{(d+1)/s} \left(\frac{1}{2}\right)^N. \quad (5.2)$$

Proof. The proof is divided into several steps.

Step I. To use the error representation (4.1) we need to bound $\Pr[W_N > \lambda]$. By Lemma 4.1 this will be done by bounding the coverage probability (4.4). We begin by discretizing the coverage problem (4.4) with an ϵ -net argument developed in [6].

Given any $\epsilon > 0$, let $\mathcal{N}_\epsilon = \{z_m\}_{m=1}^M \subset \mathbb{S}^{d-1}$ be a geodesic ϵ -net for \mathbb{S}^{d-1} with cardinality satisfying

$$M = \#(\mathcal{N}_\epsilon) \leq \left(\frac{8}{\epsilon}\right)^{d-1}.$$

Recall that the bi-cap $B_n(\lambda)$ is defined by (4.2) and (4.7) as

$$B_n(\lambda) = \text{Cap}(\varphi_n, \theta_n^+) \cup \text{Cap}(-\varphi_n, \theta_n^-).$$

Next define the shrunken bi-cap

$$T_\epsilon(B_n(\lambda)) = \text{Cap}(\varphi_n, T_\epsilon(\theta_n^+)) \cup \text{Cap}(-\varphi_n, T_\epsilon(\theta_n^-)), \quad (5.3)$$

where $T_\epsilon(\theta_n^+)$ and $T_\epsilon(\theta_n^-)$ are defined by (4.8), (4.9) and (2.7).

The key discretization step is to proceed as in [6] and note that

$$\mathbb{S}^{d-1} \not\subset \bigcup_{n=1}^N B_n(\lambda) \implies \mathcal{N}_\epsilon \not\subset \bigcup_{n=1}^N T_\epsilon(B_n(\lambda)). \quad (5.4)$$

Since the shrunken bi-caps $\{T_\epsilon(B_n(\lambda))\}_{n=1}^N$ are independent and identically distributed, (5.4) implies

$$\begin{aligned} \Pr(W_N > \lambda) &\leq \Pr\left(\mathbb{S}^{d-1} \not\subset \bigcup_{n=1}^N B_n(\lambda)\right) \\ &\leq \Pr\left(\mathcal{N}_\epsilon \not\subset \bigcup_{n=1}^N T_\epsilon(B_n(\lambda))\right) \\ &\leq \sum_{m=1}^M \Pr\left(z_m \notin \bigcup_{n=1}^N T_\epsilon(B_n(\lambda))\right) \\ &= \sum_{m=1}^M (\Pr[z_m \notin T_\epsilon(B_1(\lambda))])^N \\ &\leq \left(\frac{8}{\epsilon}\right)^{d-1} \sup_{z \in \mathbb{S}^{d-1}} (\Pr[z \notin T_\epsilon(B_1(\lambda))])^N. \end{aligned} \quad (5.5)$$

Step II. We now use (5.5) to bound $\Pr[W_N > \lambda]$ in the case when $\lambda \geq 4\delta$. In this case note that each $B_n(\lambda)$ is a genuine bi-cap that consists of two antipodal non-empty spherical caps. Since $\lambda > 2\delta$ it is straightforward from (4.2) that

$$B_n(\lambda) \supset \left\{ u \in \mathbb{S}^{d-1} : \langle u, \varphi_n \rangle > \frac{2\delta}{\lambda} \text{ or } \langle u, \varphi_n \rangle < \frac{-2\delta}{\lambda} \right\},$$

and it follows that the shrunken bi-cap $T_\epsilon(B_n(\lambda))$ satisfies

$$\begin{aligned} T_\epsilon(B_n(\lambda)) &\supset \left\{ u \in \mathbb{S}^{d-1} : |\langle u, \varphi_n \rangle| > \cos\left(\arccos\left(\frac{2\delta}{\lambda}\right) - \epsilon\right) \right\} \\ &\supset \left\{ u \in \mathbb{S}^{d-1} : |\langle u, \varphi_n \rangle| > \frac{2\delta}{\lambda} + \epsilon \right\}. \end{aligned} \quad (5.6)$$

For the remainder of this step we fix $\epsilon = \frac{2\delta}{\lambda}$. The assumption (5.1) along with (5.6) shows that if $\lambda \geq 4\delta$ then

$$\Pr[z \notin T_\epsilon(B_1(\lambda))] \leq \Pr\left(|\langle z, \varphi \rangle| \leq \frac{2\delta}{\lambda} + \epsilon\right) \quad (5.7)$$

$$\begin{aligned} &= \Pr\left(|\langle z, \varphi \rangle| \leq \frac{4\delta}{\lambda}\right) \\ &\leq \alpha \left(\frac{4\delta}{\lambda}\right)^s. \end{aligned} \quad (5.8)$$

Thus (5.5) and (5.8) imply that if $\lambda \geq 4\delta$ then

$$\Pr[W_N > \lambda] \leq \left(\frac{4\lambda}{\delta}\right)^{d-1} \left(\alpha \left(\frac{4\delta}{\lambda}\right)^s\right)^N = 4^{d-1} (4^s \alpha)^N \left(\frac{\delta}{\lambda}\right)^{sN-d+1}. \quad (5.9)$$

Since $8\delta(2\alpha)^{1/s} \geq 4\delta$ note that if $N \geq (d+2)/s$ then (5.9) implies

$$\begin{aligned} \int_{8\delta(2\alpha)^{1/s}}^{\infty} \lambda \Pr[W_N > \lambda] d\lambda &\leq 4^{d-1} (4^s \alpha)^N \int_{8\delta(2\alpha)^{1/s}}^{\infty} \lambda \left(\frac{\delta}{\lambda}\right)^{sN-d+1} d\lambda \\ &= 4^{d-1} \left(\frac{1}{2(2^s)}\right)^N \delta^2 \frac{(8(2\alpha)^{1/s})^{d+1}}{(sN-d-1)} \\ &\leq \delta^2 (32(2\alpha)^{1/s})^{d+1} \left(\frac{1}{2}\right)^N. \end{aligned} \quad (5.10)$$

Step III. Next, we bound $\Pr[W_N > \lambda]$ in the case when $0 < \lambda \leq 4(2\alpha)^{1/s}\delta$.

It will be useful to begin with the following symmetrization argument. Let $\{b_n\}_{n=1}^N$ be i.i.d. Bernoulli random variables satisfying $\Pr[b_n = 1] = \Pr[b_n = -1] = 1/2$, and additionally suppose that $\{b_n\}_{n=1}^N$ is independent of $\{\varphi_n\}_{n=1}^N$ and $\{\epsilon_n\}_{n=1}^N$. Note that the i.i.d. bi-caps $\{B_n(\lambda)\}_{n=1}^N$ have the same distribution as the i.i.d. bi-caps $\{B'_n(\lambda)\}_{n=1}^N$ defined by

$$B'_n(\lambda) = \text{Cap}(b_n \varphi_n, \theta_n^+(\lambda)) \cup \text{Cap}(-b_n \varphi_n, \theta_n^-(\lambda)).$$

This follows from (4.7), (4.8), (4.9) and the fact that ϵ_n is uniformly distributed on $[-\delta, \delta]$.

Consequently,

$$\Pr[\mathbb{S}^{d-1} \not\subset \bigcup_{n=1}^N B_n(\lambda)] = \Pr[\mathbb{S}^{d-1} \not\subset \bigcup_{n=1}^N B'_n(\lambda)].$$

Let $I(\lambda)$ denote the number of $\{\epsilon_n + \delta\}_{n=1}^N$ that lie in the interval $\left[0, \frac{\lambda}{2(2\alpha)^{1/s}}\right]$, namely:

$$I(\lambda) = \#\left(\left\{\epsilon_n + \delta\right\}_{n=1}^N \cap \left[0, \frac{\lambda}{2(2\alpha)^{1/s}}\right]\right).$$

Let

$$q(k, d-1, \alpha, s) = \Pr\left[\mathbb{S}^{d-1} \not\subset \bigcup_{n=1}^k \text{Cap}\left(b_n \varphi_n, \arccos\left(\frac{1}{2(2\alpha)^{1/s}}\right)\right)\right]. \quad (5.11)$$

Thus

$$\begin{aligned}
\Pr[W_N > \lambda] &\leq \Pr[\mathbb{S}^{d-1} \not\subset \bigcup_{n=1}^N B_n(\lambda)] \\
&= \Pr[\mathbb{S}^{d-1} \not\subset \bigcup_{n=1}^N B'_n(\lambda)] \\
&\leq \Pr[\mathbb{S}^{d-1} \not\subset \bigcup_{n=1}^N \text{Cap}(b_n \varphi_n, \theta_n^+(\lambda))] \\
&= \sum_{k=0}^N \Pr \left[\mathbb{S}^{d-1} \not\subset \bigcup_{n=1}^N \text{Cap}(b_n \varphi_n, \theta_n^+(\lambda)) \mid I(\lambda) = k \right] \Pr[I(\lambda) = k] \\
&\leq \sum_{k=0}^N \Pr \left[\mathbb{S}^{d-1} \not\subset \bigcup_{n=1}^k \text{Cap}(b_n \varphi_n, \arccos(1/2(2\alpha)^{1/s})) \right] \Pr[I(\lambda) = k] \\
&= \sum_{k=0}^N q(k, d-1, \alpha, s) \Pr[I(\lambda) = k] \\
&= \sum_{k=0}^N q(k, d-1, \alpha, s) \binom{N}{k} \left(1 - \frac{\lambda}{4\delta(2\alpha)^{1/s}}\right)^{N-k} \left(\frac{\lambda}{4\delta(2\alpha)^{1/s}}\right)^k. \tag{5.12}
\end{aligned}$$

Step IV. In this step we bound the integral $\int_0^{4\delta(2\alpha)^{1/s}} \lambda \Pr[W_N > \lambda] d\lambda$. Equation (5.12) and properties of the beta function imply

$$\begin{aligned}
&\int_0^{4\delta(2\alpha)^{1/s}} \lambda \Pr[W_N > \lambda] d\lambda \\
&\leq \sum_{k=0}^N q(k, d-1, \alpha, s) \binom{N}{k} \int_0^{4\delta(2\alpha)^{1/s}} \lambda \left(1 - \frac{\lambda}{4\delta(2\alpha)^{1/s}}\right)^{N-k} \left(\frac{\lambda}{4\delta(2\alpha)^{1/s}}\right)^k d\lambda \\
&= \sum_{k=0}^N q(k, d-1, \alpha, s) \binom{N}{k} 16\delta^2(2\alpha)^{2/s} \int_0^1 u^{k+1}(1-u)^{N-k} du \\
&= \sum_{k=0}^N q(k, d-1, \alpha, s) \binom{N}{k} 16\delta^2(2\alpha)^{2/s} \frac{(k+1)!(N-k)!}{(N+2)!} \\
&= \frac{16\delta^2(2\alpha)^{2/s}}{(N+1)(N+2)} \sum_{k=0}^N q(k, d-1, \alpha, s) (k+1) \\
&\leq \frac{16\delta^2(2\alpha)^{2/s}}{(N+1)(N+2)} \left(1 + 2 \sum_{k=1}^N k q(k, d-1, \alpha, s)\right). \tag{5.13}
\end{aligned}$$

Step V. In this step we bound the quantity $q(k, d-1, \alpha, s)$ appearing in (5.13) and defined in (5.11). We shall again employ an ϵ -net argument as in Step I. For the remainder of this

step fix

$$\epsilon = \frac{1}{2(2\alpha)^{1/s}} \left(1 - \frac{1}{4(2\alpha)^{2/s}}\right)^{-1/2}, \quad (5.14)$$

and let \mathcal{N}_ϵ be a geodesic ϵ -net for \mathbb{S}^{d-1} with cardinality $\#(\mathcal{N}_\epsilon) \leq \left(\frac{8}{\epsilon}\right)^{d-1}$. A similar argument as used to obtain (5.5) yields the following

$$\begin{aligned} q(k, d-1, \alpha, s) &= \Pr \left[\mathbb{S}^{d-1} \not\subset \bigcup_{n=1}^k \text{Cap}(b_n \varphi_n, \arccos(1/2(2\alpha)^{1/s})) \right] \\ &\leq \Pr[\mathcal{N}_\epsilon \not\subset \bigcup_{n=1}^k \text{Cap}(b_n \varphi_n, \arccos(1/2(2\alpha)^{1/s}) - \epsilon)] \\ &\leq \left(\frac{8}{\epsilon}\right)^{d-1} \left(\sup_{z \in \mathbb{S}^{d-1}} \Pr[z \notin \text{Cap}(b_1 \varphi_1, \arccos(1/2(2\alpha)^{1/s}) - \epsilon)] \right)^k. \end{aligned} \quad (5.15)$$

By (5.1) and (5.14), for an arbitrary $z \in \mathbb{S}^{d-1}$ we may compute as follows

$$\begin{aligned} &\Pr[z \notin \text{Cap}(b_1 \varphi_1, \arccos(1/2(2\alpha)^{1/s}) - \epsilon)] \\ &= \frac{1}{2} \Pr[\langle z, \varphi_1 \rangle \leq \cos(\arccos(1/2(2\alpha)^{1/s}) - \epsilon)] + \frac{1}{2} \Pr[-\langle z, \varphi_1 \rangle \leq \cos(\arccos(1/2(2\alpha)^{1/s}) - \epsilon)] \\ &= \frac{1}{2} (1 + \Pr[|\langle z, \varphi_1 \rangle| \leq \cos(\arccos(1/2(2\alpha)^{1/s}) - \epsilon)]) \\ &= \frac{1}{2} \left(1 + \Pr \left[|\langle z, \varphi_1 \rangle| \leq \frac{\cos \epsilon}{2(2\alpha)^{1/s}} + \left(1 - \frac{1}{4(2\alpha)^{2/s}}\right)^{1/2} \sin \epsilon \right] \right) \\ &\leq \frac{1}{2} \left(1 + \Pr \left[|\langle z, \varphi_1 \rangle| \leq \frac{1}{2(2\alpha)^{1/s}} + \left(1 - \frac{1}{4(2\alpha)^{2/s}}\right)^{1/2} \epsilon \right] \right) \\ &= \frac{1}{2} \left(1 + \Pr \left[|\langle z, \varphi_1 \rangle| \leq \frac{1}{(2\alpha)^{1/s}} \right] \right) \\ &\leq \frac{1}{2} \left(1 + \alpha \left| \frac{1}{(2\alpha)^{1/s}} \right|^s \right) \\ &= \left(\frac{3}{4} \right). \end{aligned} \quad (5.16)$$

Combining (5.14), (5.15), and (5.16) gives

$$\begin{aligned} q(k, d-1, \alpha, s) &\leq \left(\frac{8}{\epsilon}\right)^{d-1} \left(\frac{3}{4}\right)^k \\ &= \left(8\sqrt{4(2\alpha)^{2/s} - 1}\right)^{d-1} \left(\frac{3}{4}\right)^k \\ &\leq (16(2\alpha)^{1/s})^d \left(\frac{3}{4}\right)^k. \end{aligned} \quad (5.17)$$

Step VI. In this step we bound the sum $\sum_{k=1}^N k q(k, d-1, \alpha, s)$ appearing in (5.13). For the remainder of this step let

$$K = \frac{2d \ln(16(2\alpha)^{1/s})}{\ln(4/3)}, \quad (5.18)$$

and note that $K > 1$. If $k \geq K$ then by (5.17)

$$q(k, d-1, \alpha, s) \leq \left(\frac{3}{4}\right)^{k/2}.$$

Thus

$$\begin{aligned} \sum_{k=K+1}^N k q(k, d-1, \alpha, s) &\leq \sum_{k=K+1}^N k \left(\frac{3}{4}\right)^{k/2} \\ &\leq \sum_{k=1}^{\infty} (k+K) \left(\sqrt{\frac{3}{4}}\right)^{k+K} \\ &\leq 2K \sum_{k=1}^{\infty} k \left(\sqrt{\frac{3}{4}}\right)^k \\ &\leq 120K. \end{aligned} \quad (5.19)$$

Here we have used that

$$\forall 0 < r < 1, \quad \sum_{k=1}^{\infty} k r^k = \frac{1}{(1-r)^2}. \quad (5.20)$$

Since $0 \leq q(k, d-1, \alpha, s) \leq 1$ we also have

$$\sum_{k=1}^K k q(k, d-1, \alpha, s) \leq \sum_{k=1}^K k \leq K^2. \quad (5.21)$$

Since $K > 1$, combining (5.19) and (5.21) gives

$$1 + 2 \sum_{k=1}^N k q(k, d-1, \alpha, s) \leq 1 + 2(120K + K^2) \leq 250K^2. \quad (5.22)$$

By (5.13), (5.18), and (5.22) we conclude that

$$\begin{aligned} \int_0^{4\delta(2\alpha)^{1/s}} \lambda \Pr[W_N > \lambda] d\lambda &\leq \frac{16\delta^2(2\alpha)^{2/s}(250K^2)}{(N+1)(N+2)} \\ &\leq \frac{(16)(250)\delta^2(2\alpha)^{2/s}}{(N+1)(N+2)} \left(\frac{2d \ln(16(2\alpha)^{1/s})}{\ln(4/3)}\right)^2 \\ &\leq \frac{10^5 \delta^2 d^2 (2\alpha)^{2/s} \ln^2(16(2\alpha)^{1/s})}{(N+1)(N+2)}. \end{aligned} \quad (5.23)$$

Step VII. It remains to bound the integral $\int_{4\delta(2\alpha)^{1/s}}^{8\delta(2\alpha)^{1/s}} \lambda \Pr[W_N > \lambda] d\lambda$. Note that

$$4\delta(2\alpha)^{1/s} \leq \lambda \implies \Pr[W_N > \lambda] \leq \Pr[W_N > 4\delta(2\alpha)^{1/s}].$$

Since $4\delta(2\alpha)^{1/s} \geq 4\delta$ it follows from (5.9) that

$$\Pr[W_N > 4\delta(2\alpha)^{1/s}] \leq 4^{d-1}(4(2\alpha)^{1/s})^{d-1} \left(\frac{1}{2}\right)^N.$$

Consequently

$$\begin{aligned} \int_{4\delta(2\alpha)^{1/s}}^{8\delta(2\alpha)^{1/s}} \lambda \Pr[W_N > \lambda] d\lambda &\leq 4^{d-1}(4(2\alpha)^{1/s})^{d-1} \left(\frac{1}{2}\right)^N 24\delta^2(2\alpha)^{2/s} \\ &\leq 2\delta^2 16^d (2\alpha)^{(d+1)/s} \left(\frac{1}{2}\right)^N. \end{aligned} \quad (5.24)$$

The bound (5.2) now follows from (5.10), (5.23), and (5.24). This completes the proof. \square

Corollary 5.6. *Suppose that $\{\varphi_n\}_{n=1}^N \subset \mathbb{S}^{d-1}$ are i.i.d. versions of a unit-norm random vector φ that satisfies (5.1). There exist absolute constants $c_1, c_2 > 0$ such that if*

$$N \geq c_2 d \ln(32(2\alpha)^{1/s})$$

then the expected worst case error squared satisfies

$$\mathbb{E}|W_N|^2 \leq \frac{c_1 \delta^2 d^2 (2\alpha)^{2/s} \ln^2(16(2\alpha)^{1/s})}{(N+1)(N+2)}. \quad (5.25)$$

Example 5.7 (Uniformly distributed measurements). If $d \geq 3$ and $\{\varphi_n\}_{n=1}^N \subset \mathbb{S}^{d-1}$ are i.i.d. uniformly distributed random vectors on \mathbb{S}^{d-1} then Example 5.1 and (2.13) show that one can take $s = 1$ and $\alpha = 2C_d \leq 2\sqrt{\frac{d-1}{2\pi}} \leq \sqrt{d}$. In this case the error bound (5.25) shows that there exist constants a, c such that if $N \geq ad \ln d$ then

$$\mathbb{E}|W_N|^2 \leq \frac{c \delta^2 d^3 \ln^2 d}{(N+1)(N+2)}. \quad (5.26)$$

In the next section (see Theorem 6.1) we show that the logarithmic term in (5.26) can in fact be removed.

For perspective, note that the admissibility condition (5.1) is stronger than the condition (2.8) that was used to analyze the Rangan-Goyal algorithm. In particular, it is straightforward to show that if $\varphi \in \mathbb{S}^{d-1}$ is a random vector that satisfies (5.1), then

$$\forall x \in \mathbb{S}^{d-1}, \quad \mathbb{E}|\langle x, \varphi \rangle| \geq \left(\frac{1}{\alpha}\right)^s \left(1 - \frac{1}{s+1}\right) > 0.$$

We conclude this section by noting that Theorem 5.5 does not generally hold under the weaker condition (2.8). For this we first show if φ has a point mass then the conclusion (5.2) does not hold.

Example 5.8. Suppose $\{\varphi_n\}_{n=1}^N \subset \mathbb{R}^d$ are i.i.d. versions of a random vector φ that has a point mass at $v \in \mathbb{R}^d$ that occurs with positive probability $c = \Pr[\varphi = v] > 0$. Then

$$\forall \lambda > 0, \quad \Pr[W_N > \lambda] \geq \Pr[W_N = \infty] \geq \Pr[\varphi_1 = \varphi_2 = \dots = \varphi_N = v] = c^N.$$

Thus

$$\mathbb{E}|W_N|^2 = 2 \int_0^\infty \lambda \Pr[W_N > \lambda] d\lambda \geq 2 \int_0^\infty \lambda c^N d\lambda = \infty.$$

In particular, the conclusion (5.2) of Theorem 5.5 fails if φ has a point mass.

Example 5.9. Let $\{e_n\}_{n=1}^d \subset \mathbb{R}^d$ be an orthonormal basis. Let φ be the discrete random vector defined by $\Pr[\varphi = e_n] = 1/d$ for each $1 \leq n \leq d$. Then

$$\forall x \in \mathbb{R}^d, \quad \mathbb{E}|\langle x, \varphi \rangle| = (1/d) \sum_{n=1}^d |\langle x, e_n \rangle| \geq (1/d) \left(\sum_{n=1}^d |\langle x, e_n \rangle|^2 \right)^{1/2} = (1/d) \|x\|.$$

So, φ satisfies (2.8), but by Example 5.8, the conclusion (5.2) of Theorem 5.5 does not hold.

6. UPPER BOUNDS FOR UNIFORMLY DISTRIBUTED MEASUREMENTS

In this section we prove refined bounds for consistent reconstruction when $\{\varphi_n\}_{n=1}^N \subset \mathbb{S}^{d-1}$ are i.i.d. random vectors that are uniformly distributed on the unit-sphere. In this case, our next main result shows that the dimension dependent constant K in (1.8) is essentially dominated by the cube of the ambient dimension.

Theorem 6.1. *If the random vectors $\{\varphi_n\}_{n=1}^N \subset \mathbb{R}^d$ are i.i.d. uniformly distributed on \mathbb{S}^{d-1} and if $N \geq d + 2$ then*

$$\mathbb{E}|W_N|^2 \leq \frac{2e^{12}\delta^2 d^3}{(N+1)(N+2)} + 26\delta^2 d^{3/2} \left(\frac{11}{12} \right)^{N/2} e^{\frac{d \ln(1024d)}{2}}. \quad (6.1)$$

In particular,

$$\limsup_{d \rightarrow \infty} \mathbb{E}|W_{\lceil d^{3/2} \rceil}|^2 \leq 2e^{12}\delta^2 \quad \text{and} \quad \lim_{d \rightarrow \infty} \mathbb{E}|W_{\lceil d^{3/2} \ln d \rceil}|^2 = 0.$$

Proof. The proof is divided into several steps. We need to bound the integral

$$\mathbb{E}|W_N|^2 = 2 \int_0^\infty \lambda \Pr[W_N > \lambda] d\lambda. \quad (6.2)$$

Step I. In this step we provide preliminary bounds on $\Pr[W_N > \lambda]$ when $0 \leq \lambda \leq 2\delta\sqrt{d}$.

Note that $\{\epsilon_n + \delta\}_{n=1}^N$ are i.i.d. uniform random variables on $[0, 2\delta]$ and define the random variable

$$J(\lambda) = \# \left(\{\epsilon_n + \delta\}_{n=1}^N \cap \left[0, \frac{\lambda}{\sqrt{d}} \right] \right).$$

Lemma 4.1, (4.7), (4.8) imply that

$$\begin{aligned}
\Pr[W_N > \lambda] &\leq \Pr[\mathbb{S}^{d-1} \not\subset \bigcup_{n=1}^N B_n(\lambda)] \\
&\leq \Pr[\mathbb{S}^{d-1} \not\subset \bigcup_{n=1}^N \text{Cap}(\varphi_n, \theta_n^+(\lambda))] \\
&= \sum_{k=0}^N \Pr \left[\mathbb{S}^{d-1} \not\subset \bigcup_{n=1}^N \text{Cap}(\varphi_n, \theta_n^+(\lambda)) \mid J(\lambda) = k \right] \Pr[J(\lambda) = k] \\
&\leq \sum_{k=0}^N \Pr \left[\mathbb{S}^{d-1} \not\subset \bigcup_{n=1}^k \text{Cap}(\varphi_n, \arccos(1/\sqrt{d})) \right] \Pr[J(\lambda) = k] \\
&= \sum_{k=0}^N p(k, d-1, \arccos(1/\sqrt{d})) \Pr[J(\lambda) = k] \\
&= \sum_{k=0}^N p(k, d-1, \arccos(1/\sqrt{d})) \binom{N}{k} \left(1 - \frac{\lambda}{2\delta\sqrt{d}}\right)^{N-k} \left(\frac{\lambda}{2\delta\sqrt{d}}\right)^k \quad (6.3)
\end{aligned}$$

The above computations made use of definition (4.10) and the independence of the $\{\epsilon_n\}_{n=1}^N$ and $\{\varphi_n\}_{n=1}^N$.

Step II. In this step we bound the integral $\int_0^{2\delta\sqrt{d}} \lambda \Pr[W_N > \lambda] d\lambda$. Equation (6.3) and properties of the beta function imply

$$\begin{aligned}
&\int_0^{2\delta\sqrt{d}} \lambda \Pr[W_N > \lambda] d\lambda \\
&\leq \sum_{k=0}^N p(k, d-1, \arccos(1/\sqrt{d})) \binom{N}{k} \int_0^{2\delta\sqrt{d}} \lambda \left(1 - \frac{\lambda}{2\delta\sqrt{d}}\right)^{N-k} \left(\frac{\lambda}{2\delta\sqrt{d}}\right)^k d\lambda \\
&= \sum_{k=0}^N p(k, d-1, \arccos(1/\sqrt{d})) \binom{N}{k} (4\delta^2 d) \int_0^1 u^{k+1} (1-u)^{N-k} d\lambda \\
&= \sum_{k=0}^N p(k, d-1, \arccos(1/\sqrt{d})) \binom{N}{k} (4\delta^2 d) \frac{(k+1)!(N-k)!}{(N+2)!} \\
&= \frac{4\delta^2 d}{(N+1)(N+2)} \sum_{k=0}^N p(k, d-1, \arccos(1/\sqrt{d})) (k+1) \\
&\leq \frac{4\delta^2 d}{(N+1)(N+2)} \left(1 + 2 \sum_{k=1}^N k p(k, d-1, \arccos(1/\sqrt{d}))\right). \quad (6.4)
\end{aligned}$$

Step III. In this step we bound the sum $\sum_{k=1}^N k p(k, d-1, \arccos(1/\sqrt{d}))$ appearing in (6.4).

First note that

$$\sum_{k=1}^{60d} k p(k, d-1, \arccos(1/\sqrt{d})) \leq \sum_{k=1}^{60d} k = \frac{60d(60d+1)}{2} \leq 3600d^2. \quad (6.5)$$

Next by Lemma 4.3 and (5.20)

$$\begin{aligned} \sum_{k=60d+1}^N k p(k, d-1, \arccos(1/\sqrt{d})) &\leq \sum_{k=60d+1}^N k \left(2\sqrt{d}(13)^d \left(\frac{11}{12} \right)^{k/2} \right) \\ &= 2\sqrt{d}(13)^d \sum_{k=60d+1}^N k \left(\frac{11}{12} \right)^{k/2} \\ &\leq 2\sqrt{d}(13)^d \sum_{k=1}^{\infty} (k+60d) \left(\sqrt{\frac{11}{12}} \right)^{k+60d} \\ &\leq 2\sqrt{d}(120d) \sum_{k=1}^{\infty} k \left(\sqrt{\frac{11}{12}} \right)^k \\ &= 2\sqrt{d}(120d) \left(\frac{\sqrt{132}}{(\sqrt{12} - \sqrt{11})^2} \right) \\ &\leq 12700d^{3/2}. \end{aligned} \quad (6.6)$$

Combining (6.5) and (6.6) shows that

$$\sum_{k=1}^N k p(k, d-1, 1/\sqrt{d}) \leq 3600d^2 + 12700d^{3/2}. \quad (6.7)$$

Now (6.4) and (6.7) show that

$$\int_0^{2\delta\sqrt{d}} \lambda \Pr[W_N > \lambda] d\lambda \leq \frac{e^{12}\delta^2 d^3}{(N+1)(N+2)}. \quad (6.8)$$

Step IV. In this step we bound the integral $\int_{2\delta\sqrt{d}}^{4\delta\sqrt{d}} \lambda \Pr[W_N > \lambda] d\lambda$. Note that if $\lambda \geq 2\delta\sqrt{d}$ then (4.8) implies that $\theta_n^+(\lambda) \geq \arccos(1/\sqrt{d})$ which, in turn, implies that

$$\text{Cap}(\varphi_n, \arccos(1/\sqrt{d})) \subset \text{Cap}(\varphi_n, \theta_n^+(\lambda)) \quad (6.9)$$

Lemma 4.1 and equations (4.7) and (6.9) imply that

$$\begin{aligned}
\Pr[W_N > \lambda] &= \Pr[\mathbb{S}^{d-1} \not\subset \bigcup_{n=1}^N B_n(\lambda)] \\
&\leq \Pr[\mathbb{S}^{d-1} \not\subset \bigcup_{n=1}^N \text{Cap}(\varphi_n, \theta_n^+(\lambda))] \\
&\leq \Pr[\mathbb{S}^{d-1} \not\subset \bigcup_{n=1}^N \text{Cap}(\varphi_n, \arccos(1/\sqrt{d}))] \\
&= p(N, d-1, \arccos(1/\sqrt{d})).
\end{aligned}$$

Thus by Lemma 4.3

$$\begin{aligned}
\int_{2\delta\sqrt{d}}^{4\delta\sqrt{d}} \lambda \Pr[W_N > \lambda] d\lambda &\leq p(N, d-1, \arccos(1/\sqrt{d})) \int_{2\delta\sqrt{d}}^{4\delta\sqrt{d}} \lambda d\lambda \\
&\leq 6\delta^2 d \left(2\sqrt{d}(13)^d \left(\frac{11}{12} \right)^{N/2} \right) \\
&= 12\delta^2 d^{3/2} (13)^d \left(\frac{11}{12} \right)^{N/2}.
\end{aligned} \tag{6.10}$$

Step V. In this step we bound the integral $\int_{4\delta\sqrt{d}}^{\infty} \lambda \Pr[W_N > \lambda] d\lambda$. For this we specialize the ϵ -net argument used in the proof of Theorem 5.5 to the case when $\{\varphi_n\}_{n=1}^N$ are i.i.d. uniformly distributed on \mathbb{S}^{d-1} .

Using (5.5) and (5.7) gives

$$\begin{aligned}
\Pr[W_N > \lambda] &\leq \left(\frac{8}{\epsilon} \right)^{d-1} \sup_{z \in \mathbb{S}^{d-1}} (\Pr[z \notin T_\epsilon(B_1(\lambda))])^N \\
&\leq \left(\frac{8}{\epsilon} \right)^{d-1} \sup_{z \in \mathbb{S}^{d-1}} \left(\Pr \left[|\langle z, \varphi \rangle| \leq \frac{2\delta}{\lambda} + \epsilon \right] \right)^N.
\end{aligned} \tag{6.11}$$

Since φ is uniformly distributed on \mathbb{S}^{d-1} it follows from Example 5.1 and (2.13) that

$$\forall z \in \mathbb{S}^{d-1}, \quad \Pr \left[|\langle z, \varphi \rangle| \leq \frac{3\delta}{\lambda} \right] \leq \frac{3\delta\sqrt{d}}{\lambda}. \tag{6.12}$$

Taking $\epsilon = \delta/\lambda$ in (6.11) and using (6.12) shows that

$$\begin{aligned}
\Pr[W_N > \lambda] &\leq \left(\frac{8\lambda}{\delta} \right)^{d-1} \sup_{z \in \mathbb{S}^{d-1}} \left(\Pr \left[|\langle z, \varphi \rangle| \leq \frac{3\delta}{\lambda} \right] \right)^N \\
&\leq \left(\frac{8\lambda}{\delta} \right)^{d-1} \left(\frac{3\delta\sqrt{d}}{\lambda} \right)^N.
\end{aligned} \tag{6.13}$$

It now follows from (6.13) that if $N \geq d + 2$ then

$$\begin{aligned}
\int_{4\delta\sqrt{d}}^{\infty} \lambda \Pr[W_N > \lambda] d\lambda &\leq \int_{4\delta\sqrt{d}}^{\infty} \lambda \left(\frac{8\lambda}{\delta}\right)^{d-1} \left(\frac{3\delta\sqrt{d}}{\lambda}\right)^N d\lambda \\
&= \delta^2 (3\sqrt{d})^N 8^{d-1} \frac{(4\sqrt{d})^{d-N+1}}{N-d-1} \\
&\leq \delta^2 \sqrt{d} \left(\frac{3}{4}\right)^N (32\sqrt{d})^d.
\end{aligned} \tag{6.14}$$

Step VI. Combining (6.2), (6.8), (6.10) and (6.14) yields

$$\begin{aligned}
\frac{1}{2} \mathbb{E}|W_N|^2 &\leq \frac{e^{12}\delta^2 d^3}{(N+1)(N+2)} + 12\delta^2 d^{3/2} (13)^d \left(\frac{11}{12}\right)^{N/2} + \delta^2 \sqrt{d} \left(\frac{3}{4}\right)^N (32\sqrt{d})^d \\
&\leq \frac{e^{12}\delta^2 d^3}{(N+1)(N+2)} + 13\delta^2 d^{3/2} \left(\frac{11}{12}\right)^{N/2} (32\sqrt{d})^d \\
&= \frac{e^{12}\delta^2 d^3}{(N+1)(N+2)} + 13\delta^2 d^{3/2} \left(\frac{11}{12}\right)^{N/2} e^{\frac{d \ln(1024d)}{2}}.
\end{aligned}$$

This completes the proof. \square

Corollary 6.2. *There exist absolute constants $a, c > 0$ such that if $N \geq ad \ln d$ and $d \geq 2$ and if the random vectors $\{\varphi_n\}_{n=1}^N \subset \mathbb{R}^d$ are i.i.d. uniformly distributed on \mathbb{S}^{d-1} then*

$$\mathbb{E}|W_N|^2 \leq \frac{c d^3 \delta^2}{N^2}. \tag{6.15}$$

It is instructive to compare the dimension dependence of consistent reconstruction with linear reconstruction.

Example 6.3. Fix an error tolerance $\eta > 0$. Suppose that the random vectors $\{\varphi_n\}_{n=1}^N \subset \mathbb{S}^{d-1}$ are uniformly distributed on \mathbb{S}^{d-1} , that $\{\epsilon_n\}_{n=1}^N$ are uniformly distributed on $[-\delta, \delta]$, and that the $\{\varphi_n\}_{n=1}^N$ and $\{\epsilon_n\}_{n=1}^N$ are mutually independent.

The bound (6.15) shows that if $N \geq \frac{d^{3/2} \delta \sqrt{c}}{\eta}$ then the mean squared error for a consistent estimate \tilde{x}_{cr} satisfies

$$\mathbb{E}\|x - \tilde{x}_{\text{cr}}\|^2 \leq \mathbb{E}|W_N|^2 \leq \eta^2.$$

Thus $N = \mathcal{O}(d^{3/2})$ measurements are sufficient for consistent reconstruction to achieve η^2 -precise mean squared error.

For linear reconstruction with an arbitrary dual frame, applying (2.4) with $\sigma = \delta^2/3$ shows that linearly reconstructed \tilde{x}_{lin} satisfies

$$\mathbb{E}\|x - \tilde{x}_{\text{lin}}\|^2 \geq \frac{d^2 \delta^2}{3N}.$$

Thus at least $N \geq \frac{d^2 \delta^2}{3\eta^2}$ measurements are necessary for linear reconstruction to achieve η^2 -precise mean squared error.

7. ACKNOWLEDGMENTS

A.M. Powell was supported in part by NSF DMS 1211687 and NSF DMS 0811086, and also gratefully acknowledges the Academia Sinica Institute of Mathematics (Taipei, Taiwan) for its hospitality and support.

The authors thank Yaniv Plan, Mark Rudelson, Roman Vershynin, and Elena Yudovina for helpful comments. The authors especially thank Elena Yudovina for a suggestion which led to an improved proof of Theorem 5.5 that is more precise than an earlier version and that is also more consistent with the proof of Theorem 6.1.

REFERENCES

- [1] A. Brieden, P. Gritzmann, R. Kannan, V. Klee, L. Lovász and M. Simonivits. Deterministic and randomized polynomial-time approximation of radii. *Mathematika*, 48 (2001), 63–105.
- [2] P. Bürgisser, F. Cucker and M. Lotz. Coverage processes on spheres and condition numbers for linear programming. *The Annals of Probability*, 38 (2010), 570–604.
- [3] Z. Cvetković and M. Vetterli. On simple oversampled A/D conversion in $L^2(\mathbb{R})$. *IEEE Transactions on Information Theory*, 47 (2001), no. 1, 146–154.
- [4] Z. Cvetković. Resilience properties of redundant expansions under additive noise and quantization. *IEEE Transactions on Information Theory*, 49 (2003), no. 3, 644–656.
- [5] T.S. Ferguson. *A Course in Large Sample Theory*. Chapman & Hall, London, 1996.
- [6] L. Flatto and D.J. Newman. Random coverings. *Acta Mathematica*, 138 (1977), 241–264.
- [7] V.K. Goyal, J. Kovačević and J.A. Kelner. Quantized frame expansions with erasures. *Applied and Computational Harmonic Analysis*, 10 (2001), no. 3., 203–233.
- [8] V.K. Goyal, M. Vetterli and N.T. Thao. Quantized overcomplete expansions in \mathbb{R}^n : analysis, synthesis, and algorithms. *IEEE Transactions on Information Theory*, 44 (1998), no.1, 16–30.
- [9] L. Jacques, D.K. Hammond and J.M. Fadili. Dequantizing compressed sensing: when oversampling and non-Gaussian constraints combine. *IEEE Transactions on Information Theory*, 57 (2011), no. 1, 560–571.
- [10] S. Janson. Random coverings in several dimensions. *Acta Mathematica*, 156 (1986), 83–118.
- [11] D. Jimenez, L. Wang and Y. Wang. White noise hypothesis for uniform quantization errors. *SIAM Journal on Mathematical Analysis*, 38 (2007), no. 6, 2042–2056.
- [12] I. Jovanović and B. Beferull-Lozano. Oversampled A/D conversion and error-rate dependence of non-bandlimited signals with finite rate of innovation. *IEEE Transactions on Signal Processing*, 54 (2006), no. 6, 2140–2154.
- [13] A.M. Powell. Mean squared error bounds for the Rangan-Goyal soft thresholding algorithm. *Applied and Computational Harmonic Analysis*, 29 (2010), no. 3, 251–271.
- [14] S. Rangan and V.K. Goyal. Recursive consistent estimation with bounded noise. *IEEE Transactions on Information Theory*, 47 (2001), no. 1, 457–464.
- [15] A.F. Siegel and L. Holst. Covering the circle with random arcs of random sizes. *Journal of Applied Probability*, 19 (1982), no. 2, 373–381.
- [16] H. Solomon. *Geometric Probability*. Society for Industrial and Applied Mathematics, Philadelphia, 1978.
- [17] A.J. Stam. Limit theorems for uniform distributions on spheres in high-dimensional Euclidean spaces. *Journal of Applied Probability*, 19 (1982), no. 1, 221–228.
- [18] N.T. Thao and M. Vetterli. Deterministic analysis of oversampled A/D conversion and decoding improvement based on consistent estimates. *IEEE Transactions on Signal Processing*, 42 (1994), no. 3, 519–531.
- [19] N.T. Thao and M. Vetterli. Reduction of the MSE in R -times oversampled A/D Conversion from $\mathcal{O}(R^{-1})$ to $\mathcal{O}(R^{-2})$. *IEEE Transactions on Signal Processing*, 42 (1994), 200–203.
- [20] N.T. Thao and M. Vetterli. Lower bound on the MSE in over sampled quantization of periodic signals using vector quantization analysis. *IEEE Transactions on Information Theory*, 42 (1996), 469–479.
- [21] J.G. Wendel. Note on the gamma function. *American Mathematical Monthly*, 55 (1946), no. 9, 563–564.

VANDERBILT UNIVERSITY, DEPARTMENT OF MATHEMATICS, NASHVILLE, TN 37240, USA

E-mail address: alexander.m.powell@vanderbilt.edu

E-mail address: tyler.whitehouse@gmail.com